

CANADIAN JOURNAL OF PHILOSOPHY, 2017  
VOL. 47, NO. 1, 116–147  
<http://dx.doi.org/10.1080/00455091.2016.1235856>

 **Routledge**  
Taylor & Francis Group

## Kant's deductions of morality and freedom

Owen Ware

Department of Philosophy, Simon Fraser University, Burnaby, Canada

### ABSTRACT

It is commonly held that Kant ventured to derive morality from freedom in *Groundwork* III. It is also believed that he reversed this strategy in the second *Critique*, attempting to derive freedom from morality instead. In this paper, I set out to challenge these familiar assumptions: Kant's argument in *Groundwork* III rests on a moral conception of the intelligible world, one that plays a similar role as the 'fact of reason' in the second *Critique*. Accordingly, I argue, there is no reversal in the proof-structure of Kant's two works.

**ARTICLE HISTORY** Received 18 June 2015; Accepted 9 September 2016

**KEYWORDS** Kant; morality; freedom; justification; deduction; fact of reason

### 1. Introduction

There is little consensus among scholars over how to interpret Kant's deductions of morality and freedom in his mature ethics – with one striking exception. Nearly every commentator believes that Kant rejected his argument from *Groundwork* III (1785) soon after its publication.<sup>1</sup> Aside from their disagreements, scholars have come to unite on this basic point of interpretation.<sup>2</sup>

The topic of this paper is whether or not this interpretation is true. Despite the fact that Kant never acknowledged a change in the structure of his arguments, the standard view in the literature is that his doctrine of the 'Fact of Reason' from the second *Critique* (1788) marks a turn in his thinking. For rather than establish our freedom independently – either on the grounds that we must engage in deliberation or on the grounds that theoretical reason is self-active within us – Kant now treats our consciousness of the moral law's authority as an inexplicable 'fact,' one that admits of no deduction but that secures the reality of freedom in beings 'who cognize this law as binding upon them' (KpV, AA 5:47).<sup>3</sup> In an unexpected move, Kant appears to reverse his strategy of justification, working from freedom to morality in 1785 and then from morality to freedom in 1788.<sup>4</sup> In what follows, I will argue that this widely accepted reading of Kant's

**CONTACT** Owen Ware  [oware@sfu.ca](mailto:oware@sfu.ca)

© 2016 Canadian Journal of Philosophy

deductions, what we may call the Reversal Reading, is mistaken. In its place, I will defend a Continuity Reading.<sup>5</sup>

Admittedly, there is only so much I can hope to accomplish in the following pages. What I plan to offer is a *sketch* of an alternative interpretation, the details of which would take us well beyond the scope of a single paper. I say this now in order to forestall worries readers might have about the completeness of my proposal. As much as I will try to cover all the key points of dispute over the relationship between *Groundwork* III and the second *Critique*, the fact remains that we lack sufficient space to satisfy every point of interest. My aim is not so much to settle the dispute once and for all, but to make the Continuity Reading a plausible (though perhaps unorthodox) candidate for understanding Kant's project of moral justification. I shall be content, then, merely to shift the burden of proof: by the end of this paper, I hope it is clear that proponents of the Reversal Reading must defend, rather than assume, the truth of their position. Whatever the outcome, I believe the current literature stands in need of such re-examination.

I begin (in Section 2) by reviewing the early reception of Kant's ethics in the late eighteenth and early nineteenth centuries. A curious historical point is that Kant's first interpreters saw no fundamental difference between the proof-structure of *Groundwork* III and the second *Critique*. In fact, the Reversal Reading did not acquire uniform shape in the literature until 1960, the year Dieter Henrich and Lewis White Beck published their studies on Kant's project of justification in ethics. With this background in place, I turn attention to the details of *Groundwork* III itself (in Section 3), showing that contemporary readers have found the Reversal Reading attractive because it purports to explain Kant's worry about a 'hidden circle.' I then proceed to outline an alternative reading of this circle, one that is consistent with Kant's argument from morality to freedom in the second *Critique*. My full defense of the Continuity Reading comes next (in Section 4), and I devote a further section to answering potential objections to it (in Section 5). Then, after explaining why Kant's two works still differ in their presentation (in Section 6), I close by briefly considering what we can learn from Kant's project of justification today (in Section 7).

## 2. Historical background

Those familiar with the current literature will no doubt be surprised to learn that Kant's first interpreters saw no change in his mature ethical works. In the decades to follow the publication of the second *Critique*, commentators embraced the doctrine of the 'Fact of Reason' (*Faktum der Vernunft*) as a paradigm for philosophical inquiry, not only in the domain of morality, but also – more controversially – in the domain of knowledge. As early as 1786/87, Karl Reinhold, a devoted advocate of Kant's philosophy, introduced the concept of 'facts of consciousness'<sup>6</sup> (*Tatsachen des Bewußtseins*) – a phrase that was quickly picked

up by Johannes Gottlieb Fichte in his formulation of a ‘science of knowledge’ (1794/95). Neither of these authors spoke of Kant’s *Faktum* as a novel development within his system of ethics. Nor was this unusual for the time. In his book-length commentary on the second *Critique* (1796–97), for example, Christian Friedrich Michaelis discussed why our consciousness of the moral law is a ‘fact’ (in the sense of a *Tatsache*) that admits of no further proof, a point he contrasted more than once to Kant’s transcendental deduction of the categories in the first *Critique*.<sup>7</sup> In this two-volume work, amounting to over 600 pages in length, Michaelis only mentioned Kant’s *Groundwork* five times, and each time he did so without reference to the Fact of Reason.<sup>8</sup>

This is not to say Kant’s arguments from *Groundwork* III and the second *Critique* won the immediate approval of his contemporaries. Many interpreters of these books were of a critical bent.<sup>9</sup> Yet what is interesting is that none of these early writers even suggested that Kant’s later doctrine of the Fact of Reason marked a change, let alone a reversal, in his project of justification. Looking at them now, one gets the impression that commentators were either uninterested in the question of Kant’s intellectual development from 1785 to 1788, or they were willing to accept his word that the second *Critique* ‘presupposes’ the *Groundwork* (KpV, AA 5:8). Indeed, this is how Georg Albert Mellin organized the relationship between these texts in his monumental *Encyclopedic Dictionary of Critical Philosophy* at the turn of the nineteenth century. After summarizing the three sections of the *Groundwork*, Mellin paraphrased Kant’s remark that the second *Critique* presupposes the *Groundwork* insofar as it ‘constitutes preliminary acquaintance with the principle of duty and provides and justifies a determinate formula of it’ (KpV, AA 5:8).<sup>10</sup> Like Reinhold, Fichte, and Michaelis before him, Mellin also spoke of the Fact of Reason as a kind of *Tatsache* without ever mentioning Kant’s earlier deductions. Not a single sentence in the *Encyclopedic Dictionary*’s six volumes hinted at anything like a rupture between *Groundwork* III and the second *Critique*.

All of this stands in sharp contrast to the Reversal Reading popular today. So we must ask: When did commentators come to believe that Kant’s project of justification underwent a turn? The answer is elusive in part because attitudes toward Kantian philosophy in the early nineteenth century were inconsistent. Not everyone was as enthusiastic as Reinhold and Fichte, for instance. Still, what we can say for certain is that by the early 1800s, 20 years after the publication of the second *Critique*, the positive reception of Kant’s ethics was on the wane. There was a growing cynicism among scholars of the time, not only toward the system of Kantian philosophy, but also toward the specific doctrine Kant’s early advocates cherished: the doctrine of the Fact of Reason. In his book *On the Spirit of Philosophy* (1803), for example, Jakob Salat complained that Kant’s appeal to a *Faktum* reduced the foundations of his practical philosophy to ‘mysticism’ – anticipating Hegel’s famous remark that ‘cold duty is the final undigested lump left within the stomach, the revelation given to reason.’<sup>11</sup> If there was a

shared sentiment among scholars, it was that the doctrine of the Fact of Reason committed a gross error, although there was no consensus at the time over the exact nature of this error.

Surprisingly, the Reversal Reading did not acquire uniform shape in the literature until 1960, the year Dieter Henrich published his essay, 'Der Begriff der sittlichen Einsicht und Kants Lehre vom Faktum der Vernunft,'<sup>12</sup> and Lewis White Beck published his book, *A Commentary on Kant's Critique of Practical Reason*. Each author reached the same verdict on Kant's deductions. As Beck put it, the argument of the second *Critique* 'takes a truly astonishing turn' (1960, 173). By having said 'that the principle [of morality] needs no deduction, he apparently stands the argument of the *Groundwork* on its head. He uses the moral law, the Fact of Reason, as the *prius* to deduce something else, namely, freedom, which is its *ratio essendi*' (1960, 172).<sup>13</sup> Henrich went further to explain the cause of this apparent reversal, suggesting that the argument of the *Groundwork* was still driven by Kant's pre-critical project of deriving morality from theoretical reason. 'The difference in the structure of the *Groundwork* and the *Critique of Practical Reason* can be understood,' he wrote, 'only if we realize that Kant did not yet see clearly in the *Groundwork* that a deductive justification of ethics must necessarily turn out to be unsatisfactory and contradictory.'<sup>14</sup> What is strange, given the reception of Kant's ethics I have just reviewed, is that the opinion Beck and Henrich expressed has gone largely unchallenged by scholars over the last 50 years.<sup>15</sup>

### 3. The argument of *Groundwork* III

I believe many contemporary readers have found the Reversal Reading attractive because it purports to explain Kant's worry about a 'hidden circle' (*geheimer Zirkel*) in *Groundwork* III. In what follows, I will outline an alternative reading of this circle, one that is consistent with Kant's argument from morality to freedom in the second *Critique* (Section 3.1).<sup>16</sup> I will then argue that Kant's concept of an 'intelligible world' (*intelligiblen Welt*) provides the key to his deduction of freedom, but only if we interpret this concept as a moral ideal (Section 3.2). As we shall see, one virtue of this reading is that it makes sense of Kant's opening remark in *Groundwork* III that he plans to offer a 'deduction of the concept of freedom from pure practical reason, and with it the possibility of a categorical imperative' (GMS, AA 4:447; my emphasis).<sup>17</sup>

#### 3.1. A 'hidden circle'

Kant's worry about a circle in *Groundwork* III is a bit mystifying, since he says little to clarify what the problem amounts to. At first, all he says is that 'one must freely confess that a kind of circle appears here from which, as it seems, there is no way out' (GMS, AA 4:450). Now if we assume that the threat of a

circle emerges from using a moral premise to establish our freedom, then we can agree that our only means to escape the circle would be to find *non-moral* evidence for thinking we are free. On this assumption, it would be natural to hear Kant's remarks in the second *Critique* as departing from his earlier approach, as when he says that the moral law, while admitting of no deduction, 'conversely itself serves as the principle of the deduction of an inscrutable faculty,' that is, the 'faculty of freedom' (KpV, AA 5:47). Yet the naturalness of this reading rests on how we characterize Kant's worry in *Groundwork* III. If the threat of a circle does not come from using a moral premise to establish our freedom, then it is far from obvious that our only means of escape is to find a non-moral route.

Let us consider what Kant says more closely:

One must freely confess that a kind of circle appears here from which, as it seems, there is no way out. We take ourselves as free in the order of efficient causes in order to think of ourselves under moral laws in the order of ends, and we afterwards think of ourselves as subject to these laws because we have attributed to ourselves freedom of will. (GMS, AA 4:450)

Upon reading this, it seems Kant is saying that a circle arises if we proceed from a moral starting point – for example, if we consider ourselves under moral laws in the order of ends. Consequently, it is tempting to think that our only hope to escape the circle lies in finding a non-moral premise, one that would establish the validity of moral laws without begging the question. While the content of this premise is a topic of much controversy, the majority of scholars agree that Kant's argument has this proof-structure. They believe we should find non-moral evidence for thinking we are free, either on the grounds that we must engage in deliberation or on the grounds that theoretical reason is self-active within us (see Section 5.2 below).

While tempting, I find this familiar characterization of Kant's worry problematic. When we look at how he describes the circle, it is by no means clear that our only hope of escape lies in finding a non-moral route. As we read further, Kant suggests that our only hope of escape lies in finding a non-analytic or 'synthetic' route. Consider what he says, for instance, in the remainder of the passage quoted above:

[F]reedom and the will's own lawgiving are both autonomy, hence reciprocal concepts, yet for this very reason *one cannot be used to explain the other or to specify its ground*, but can at most be used only for the logical aim of bringing different representations of the same object to a single concept (as different fractions of the same value are brought to their lowest expression). (GMS, AA 4:450; my emphasis)

The problem of a circle emerges, then, because the concepts of freedom, autonomy, and morality are all co-entailing. Analyzing freedom leads us to the concept of autonomy, and analyzing autonomy leads us to the concept of morality (GMS, AA 4:447). Yet, the movement from one to another is not a movement of justification: it is a movement of clarification. At no point have we established a 'ground' (*Grund*) for anything. All we have shown, rather, is that apparently

different representations reduce to a common source. And that is true regardless of our starting point. Whether we start from freedom and end up with morality, or whether we start from morality and end up with freedom, we cannot *conceptualize our way* to the validity of a concept.

This begins to explain why Kant voices his worry in the first place. What commentators often overlook is that Kant is trying to motivate the final ‘transition’ (*Übergang*) of the *Groundwork*, the transition from a ‘metaphysics of morals’ (the task of *Groundwork* II) to a ‘critique of pure practical reason’ or what he also calls a ‘critique of the subject’ (the task of *Groundwork* III) (GMS, AA 4:440).<sup>18</sup> In Kant’s view, a metaphysics of morals is effective for investigating our commonly received notions of duty. When we consider the structure of imperatives and their relation to rational beings in general, we can specify the formula of the moral law in terms of autonomy. We would never have been able to see this connection had we – like the popular philosophers of Kant’s day – adopted a mixed method of inquiry. And yet, as we learn by the end of *Groundwork* II, the method of a metaphysics of morals is also limited. Since it does nothing more than analyze and clarify concepts, it is unfit to justify our everyday notions of duty (GMS, AA 4:445). What we need, Kant thinks, is a new method, what he calls a ‘critique.’ Only a critique – that is, an investigation into the bounds of our faculty of reason – will decide whether we are justified in thinking we are free.

My suggestion is that Kant is advancing a methodological point at the beginning of *Groundwork* III. When we employ the method of a metaphysics of morals, we are entitled to make analytic claims about freedom, autonomy, and morality since we are only attempting to explicate their conceptual structure. The moment we actually claim to be free, however, we are no longer within the sphere of analysis: we are asserting a synthetic claim of real possibility, and such a claim now demands justification. Without this justification, our claim to be free would lack proper warrant – in short, we would be begging the question. Yet, my point is that we would be begging the question either way: so long as we only employ the method of a metaphysics of morals, we have no basis to assert the reality of freedom or the validity of the categorical imperative. So Kant’s worry about a circle is his way of telling us that we need to go beyond the sphere of analysis and examine our faculty of reason instead, but it is left open at this point whether he thinks the ultimate proof of our freedom will be moral or non-moral in character.

### **3.2. *The intelligible world: a moral ideal***

After introducing the problem of a circle, Kant tells us there is only ‘one way out, namely to inquire whether we do not take a different standpoint when by means of freedom we think ourselves as causes efficient a priori than when we represent ourselves in terms of our actions as effects that we see before our eyes’ (GMS, AA 4:450). This different standpoint leads him to the concept of an



intelligible world, which we arrive at by abstracting from every condition of the world of sense, including our needs, impulses, and inclinations. What is left over from this abstraction is an idea, the idea of a world governed by rational laws. But Kant is clear that the force of the idea is entirely practical: it gives us nothing to intuit or cognize.<sup>19</sup> Thus, in a revealing passage near the end of *Groundwork* III, he writes that the idea of an intelligible world

signifies only a 'something' that is left over when I have excluded from the determining grounds of my will everything belonging to the world of sense, merely in order to limit the principle of motives from the field of sensibility by circumscribing this field and showing that it does not include everything within itself but that there is still more beyond it; but of this something more I have no further cognizance. As for pure reason, which thinks this ideal, after its isolation from all matter, that is, cognition of objects, nothing is left for me but the form of it – namely the practical law of the universal validity of maxims – and to think of reason, conformably with this, with reference to a pure world of understanding as a possible efficient cause, that is, a cause determining the will. (GMS, AA 4:462)<sup>20</sup>

The general direction of Kant's argument comes out clearly here. He is saying we can think of the law of an intelligible world – as the 'practical law of the universal validity of maxims' – and we can think of our reason in conformity with this law – as a 'cause determining the will.' In doing so, we recognize our capacity to act independently of all motives coming from the world of sense. We recognize that our faculty of reason has a pure use beyond the satisfaction of our needs, impulses, and inclinations. As the cited passage also makes clear, Kant is saying that when we think of ourselves in conformity with the law of an intelligible world, we have grounds to ascribe freedom to ourselves. From the standpoint of this ideal, we can go beyond mere analysis and claim to be free, but – and this is the crucial point – the basis of our claim is of a moral character, for it rests on our consciousness of a pure practical law.<sup>21</sup>

Support for this reading also comes from what Kant says about our common presupposition of freedom earlier in the text. The 'legitimate claim even of common human reason to freedom of the will,' he writes, 'is founded on the consciousness and the granted presupposition of the independence of reason from merely subjectively determining causes' (GMS, AA 4:457). In the very next paragraph, Kant links this consciousness of independence to the 'law' we give ourselves in the intelligible world:

So it is that the human being claims for himself a will which lets nothing be put to his account that belongs merely to his desires and inclinations, and on the contrary thinks as possible by means of it, indeed as necessary, actions that can be done only by disregarding all desires and sensible incitements. The causality of such actions lies in him as intelligence and in the laws of effects and actions in accordance with principles of an intelligible world, of which he knows nothing more than that in it reason alone, and indeed pure reason independent of sensibility, gives the law. (GMS, AA 4:457)<sup>22</sup>

Our warrant for thinking we are free therefore comes not merely from our membership in an intelligible world, but from our consciousness of the law we give to ourselves in this world – the moral law. No longer limited to the world of sense and the practical perspective it affords us, then, we can at last see the real possibility of determining our will to action on the basis of pure reason alone. And since we are no longer analyzing a relationship between concepts, the threat Kant used to motivate his discussion, the threat of a ‘hidden circle,’ is now out of place.

Of course, one might still wonder: Even if we are entitled to think of ourselves as free for the reasons just spelled out, what *right* do we have to speak of those moral laws governing the intelligible world? Do they not stand in need of justification as well? Kant’s answer is clear, but it only becomes explicit at the end of *Groundwork* III. Given the nature of human reason, he explains, we can only comprehend what happens or what ought to happen on the basis of prior conditions (GMS, AA 4:463). Yet the laws of morality convey absolute necessity. So when it comes to those laws themselves, the most Kant says we can do is appreciate the fact that our philosophical investigation has come to an end. That means we cannot secure a deduction of morality proper, but we can at least see why this is no fault of our own. ‘It is,’ he writes, ‘no criticism of our deduction of the supreme principle of morality, but an accusation that would have to be brought against human reason in general, that it cannot make comprehensible, as regards its absolute necessity, an unconditional practical law (such as the categorical imperative must be)’ (GMS, AA 4:463).<sup>23</sup>

Bringing these pieces together, we can now reconstruct Kant’s argument into the following steps:

- Step 1* When we think of ourselves as ‘causes efficient a priori,’ we presuppose our membership in a world beyond the influences of our sensible nature.
- Step 2* This ‘merely intelligible’ world gives us nothing to intuit or cognize. But it gives us an ideal for thinking of our will in conformity with pure practical laws.
- Step 3* When we think of our will in this way, we have a standpoint for cognizing our freedom: our capacity to act by reason alone. This capacity is equivalent to autonomy.

Having reached this point in his discussion, Kant is able to connect his deduction back to the formula of morality he had established earlier in *Groundwork* II:

- Step 4* Autonomy is equivalent to the formula of morality: ‘so act as if your maxims were to serve at the same time as a universal law (for all rational beings).’

With this outline in hand, we are ready to turn to the second *Critique*. But first, let us recall that Kant had also promised us a second argument, writing that he would



offer a 'deduction of the concept of freedom from pure practical reason, and *with it* the possibility of a categorical imperative' (GMS, AA 4:447; my emphasis). After avoiding the threat of a circle, Kant goes on to say that the moral law takes the form of an imperative for us because it has its source within us: what we 'would' do were we only members of an intelligible world is what we 'ought' to do as beings who belong to the world of sense 'at the same time' (GMS, AA 4:454).<sup>24</sup> While unpacking this argument would take us beyond the scope of this paper, I mention it here only to make the following point salient: in *Groundwork* III, Kant never equates a deduction of the categorical imperative with a deduction of the moral law itself. As he makes clear, the latter is beyond the comprehension of human reason and so beyond the reach of *any* philosophical proof.<sup>25</sup>

## 4. The argument of the second *Critique*

### 4.1. A 'vainly sought deduction'

Those sympathetic to the Reversal Reading may now wonder how I can explain Kant's apparent confession in the second *Critique* that his earlier strategy of justification was a failure. In a much-discussed passage, Kant writes that 'the moral law cannot be proved by any deduction,' adding:

But something different and quite paradoxical takes the place of this vainly sought deduction of the moral principle, namely that the moral principle conversely itself serves as the principle of the deduction of an inscrutable faculty which no experience could prove but which speculative reason had to assume as at least possible [ ... ] namely the faculty of freedom, of which the moral law, which itself has no need of justifying grounds, proves not only the possibility but the reality in beings who cognize this law as binding upon them. (KpV, AA 5:47)

As far as confessions go, this passage seems to present us with strong evidence that Kant changed his mind by the time he composed the second *Critique*. For what could this 'vainly sought deduction of the moral principle' refer to, if not the very deduction Kant had ventured in 1785?<sup>26</sup>

This passage acquires a different meaning, I believe, when we read it alongside the 'concluding remark' of *Groundwork* III just mentioned. In this final paragraph, Kant explains why our faculty of reason faces a dilemma (GMS, AA 4:463). On the one hand, our reason is driven to seek the condition of what happens or what ought to happen with the aim of finding something absolutely necessary, something that would end the regress of conditions. On the other hand, due to the very limitations of reason itself, we can only comprehend what happens or what ought to happen when we discover its underlying condition. So the very principle driving the activity of reason itself, namely, to find something unconditionally necessary, can never be satisfied. As Kant puts it, reason 'restlessly seeks the unconditionally necessary and sees itself constrained to assume it without any means of making it comprehensible to itself, fortunate enough if

it can discover only the concept that is compatible with this presupposition; that is, the concept of freedom (GMS, AA 4:463).

This is why Kant thinks we should not hope to explain the moral law's necessity, such as a deduction would require of us, because its necessity is by definition unconditional. In other words, a deduction of the moral law would be vainly sought since we cannot 'comprehend the practical unconditional necessity of the moral imperative' (GMS, AA 4:463). When Kant makes this point explicit in the second *Critique*, he may have *Groundwork* III in mind, but not because he is repudiating his earlier approach. On my reading, Kant may be wanting to draw the same lesson in each text, that human insight reaches an end as soon as we are dealing with fundamental laws or fundamental faculties (KpV, AA 5:47). That is why he thinks seeking to comprehend the moral law would be futile. All we can do is comprehend its 'incomprehensibility' (*Unbegreiflichkeit*), and that is, Kant says, 'all that can be justly required of a philosophy that strives in its principles to the boundary of human reason' (GMS, AA 4:463). On this point at least *Groundwork* III and the second *Critique* stand in agreement: a deduction of the moral law is beyond our explanatory reach.<sup>27</sup>

Granted, when Kant says that the moral law 'cannot be proved by any deduction,' it is hard to rid ourselves of the impression that he is referring to his earlier strategy of argument. Yet further study reveals that this impression is deceiving since the argument of *Groundwork* III refers to a deduction of the categorical imperative, *not* to a deduction of the moral law (GMS, AA 4:454). The insight Kant is seeking in the final section of the *Groundwork* is insight into how the moral law takes the form of an imperative (or 'ought') for finite rational beings like us. Strictly speaking, then, he is seeking a deduction of the law's 'necessitation' (*Nötigung*), not a deduction of its 'necessity' (*Notwendigkeit*). As we have seen, Kant denies that we could ever attain the latter, given the subjective limitations of our reason (the fact that we only comprehend things on the basis of prior conditions). As a result, Kant's remark about a 'vainly sought deduction' does not support the Reversal Reading since the deduction he affirms in *Groundwork* III and the deduction he denies in the second *Critique* are not one and the same.

Interestingly, the only place where Kant makes a direct reference to the *Groundwork* is in the Preface to the second *Critique*, but what he says there is seemingly positive:

I must leave it to connoisseurs of a work of this kind to estimate whether such a system of pure practical reason as is here developed from the critique of it has cost much or little trouble, especially so as not to miss the right point of view from which the whole can be correctly traced out. It presupposes, indeed, the *Groundwork for the Metaphysics of Morals*, but only insofar as this constitutes preliminary acquaintance with the principle of duty and provides and justifies [*rechtfertigt*] a determinate formula of it; otherwise, it stands on its own. (KpV, AA 5:8)

For decades after the publication of the second *Critique*, readers would refer to this passage to explain the relationship between Kant's two works, as we saw

with Mellin's *Encyclopedic Dictionary*. Looking at this passage now, there is no trace of a rupture in Kant's conception of the path he took from 1785 to 1788. Instead of confessing a change of mind, he underlines the continuity of his project, not only with his formulation of the moral principle (the task of *Groundwork* I and II), but also with his justification of it (the task of *Groundwork* III). What is more, the language of KpV, AA 5:8 is strongly reminiscent of what Kant says in the Preface to the *Groundwork*: that the text is 'nothing more than the search for and establishment of' the principle of morality (GMS, AA 4:392). When Kant later says that the second *Critique* 'presupposes' the *Groundwork*, he may not be restricting his statement to the task of clarification, for in each place, he speaks of justification ('*Rechtfertigung*' or '*Festsetzung*').<sup>28</sup> In any case, there is no direct evidence here to support the Reversal Reading.<sup>29</sup>

#### 4.2. *The continuity reading*

All the pieces for my alternative account are now in place. To begin with, it is clear that Kant still accepts the first step of his 1785 deduction:

*Step 1* When we think of ourselves as 'causes efficient a priori,' we presuppose our membership in a world beyond the influences of our sensible nature.

One difference is that in the second *Critique*, Kant does not move directly to:

*Step 2* This 'merely intelligible' world gives us nothing to intuit or cognize. But it gives us an ideal for thinking of our will in conformity with pure practical laws.

Instead, Kant appeals to the Fact of Reason: 'It is therefore the moral law,' he writes, 'that *first* offers itself to us and, inasmuch as reason presents it as a determining ground not to be outweighed by any sensible conditions and indeed quite independent of them, leads directly to the concept of freedom' (KpV, AA 5:29–30; cf., 5:4n). Nevertheless, it is clear that Step 2 of *Groundwork* III plays a similar role as the Fact of Reason, for in both cases, Kant is appealing to our consciousness of pure practical laws to establish our freedom. So both the arguments of *Groundwork* III and the second *Critique* use a moral premise to infer:

*Step 3* When we think of our will in this way, we have a standpoint for cognizing our freedom: our capacity to act by reason alone. This capacity is equivalent to autonomy.

Nor do the parallels between Kant's two works end at this point. Later, in the second *Critique*, he explains that the law of autonomy 'is the moral law, which is therefore the fundamental law of a supersensible nature and of a pure world of the understanding' (KpV, AA 5:42). Here, it is worth recalling that Kant made

a similar move in *Groundwork* III when he connected the deduction of freedom back to his formula of the moral principle,<sup>30</sup> leading us to:

*Step 4* Autonomy is equivalent to the formula of morality: ‘so act as if your maxims were to serve at the same time as a universal law (for all rational beings).’

Finally, while the question of how a moral ‘ought’ is possible is not central to the second *Critique*, Kant does make a point of addressing it. In the case of human beings, he explains, ‘the law has the form of an imperative, because in them, as rational beings, one can presuppose a *pure* will but, insofar as they are beings affected by needs and sensible motives, not a *holy* will, that is, such a will as would not be capable of any maxim conflicting with the moral law’ (KpV, AA 5:32).<sup>31</sup> Just as we saw in *Groundwork* III, Kant is saying that moral imperatives are possible for us because they have their source within us: what we ‘would’ do were we only members of an intelligible world is what we ‘ought’ to do as beings who belong to the world of sense ‘at the same time’ (GMS, AA 4:454). On the whole, then, there is remarkable consistency in the layout of each argument. Kant is not venturing a deduction of the moral law itself, but in both places, he is offering a deduction of freedom, followed by an account of how the moral law takes the form of an imperative.<sup>32</sup>

## 5. Objections and replies

Stepping back, there are at least three objections one could raise against the preceding sections of this paper. One could argue that I have overlooked Kant’s solution to the circle problem in *Groundwork* III (Section 5.1), and one could argue that I have overlooked evidence that Kant was seeking a non-moral route to freedom, both in the *Groundwork* (Section 5.2) and in his earlier writings (Section 5.3).

### 5.1. First objection: the circle problem (again)

Kant in fact gives us two characterizations of the circle problem in *Groundwork* III, and commentators often draw upon the second to support the view that he was seeking a non-moral argument. On this occasion, Kant writes:

The suspicion that we raised above is now removed, the suspicion that a hidden circle was contained in our inference from freedom to autonomy and from the latter to the moral law – namely that we perhaps took as a ground the idea of freedom only for the sake of the moral law, so that we could afterwards infer the latter in turn from freedom, and that we were thus unable to furnish any ground at all for the moral law but could put it forward only as a *petitio principii* [or ‘question-begging’ claim] disposed souls would gladly grant us, but never as a demonstrable proposition. (GMS, AA 4:453)

Here, one might worry that if my reading is correct, and Kant was appealing to a moral premise to establish our freedom of will, then he should be saying something else in this passage – that the suspicion of a circle has been confirmed. On my reading, it may seem that Kant was starting with the moral law (something ‘disposed souls would gladly grant us’) in a way that would be question-begging if directed to a skeptic (a ‘*petitio principii* [ ... ] but never a demonstrable proposition’). When we return to the cited passage, however, Kant is clear in saying that he *has* established the moral law as a demonstrable proposition, and this suggests he was arguing to, rather than from, a moral claim, just as the standard reading tells us.<sup>33</sup>

As compelling as this evidence appears, let us first consider the meaning behind Kant’s terminology more closely. To be sure, talk of a ‘demonstrable proposition’ seems to be precisely what Kant denies in the second *Critique*: that we can grasp the moral law’s necessity as a principle for all rational beings. And if this is what he has in mind in *Groundwork* III, there would be good reason to think his project of justification changes significantly in later works. Yet on reflection, it turns out that Kant’s reference at GMS, AA 4:453 is open to interpretation. For example, the proposition he says requires demonstration could refer to his formula of the moral law, rather than to the law’s supreme normative authority (which Kant later says we *cannot* comprehend).<sup>34</sup> This would be consistent with my reading because the ‘moral premise’ I have claimed Kant employs to establish our freedom of will (the pure rational law we cognize from the standpoint of an intelligible world in Step 2) is distinct from the ‘moral conclusion’ he works toward (the formula of this law in terms of a principle of autonomy in Step 4). In view of this distinction, there is no conflict between the account I favor and Kant’s second characterization of the circle problem. There are other ways to understand his talk of a ‘demonstrable proposition,’ and for the purposes of this paper, that all I need to make the Continuity Reading a plausible alternative.

## 5.2. Second objection: a non-moral route in the *Groundwork*

A second objection one could raise is that my account of *Groundwork* III overlooks two passages where Kant appears to be offering non-moral evidence for our freedom of will. The first concerns his remark that rational beings must act under the idea of freedom (Section 5.2.1); the second concerns his remark that theoretical reason is self-active within us (Section 5.2.2).

### 5.2.1. First passage: the idea of freedom

Starting with the first passage, Kant writes:

I say now: every being that cannot act otherwise than *under the idea of freedom* is just because of that really free in a practical respect, that is, all laws that are inseparably bound up with freedom hold for him just as if his will had been validly pronounced free also in itself and in theoretical philosophy. (GMS, AA 4:448)<sup>35</sup>

Given its placement in the text, it may look like Kant is advancing the following argument: (i) that freedom and morality are co-entailing concepts, (ii) that we must think of ourselves under the idea of freedom, and (iii) that morality is therefore valid for us. The problem with this interpretation is that Kant does not intend his remark about acting ‘under the idea of freedom’ as a premise in his deduction. The remark occurs in subsection two where Kant is offering nothing more than ‘preparatory elucidation’ (GMS, AA 4:447). Afterwards, he explains that all he has done in subsection two is trace the ‘determinate concept of morality back to the idea of freedom,’ indicating that his discussion belongs to the analytic side of *Groundwork* III, that is, before the transition to a critique. Kant also remarks that ‘we could not, however, prove freedom as something actual even in ourselves or in human nature; we saw only that we must presuppose it if we want to think of a being as rational and endowed with consciousness of its causality with regard to its actions, i.e. with a will’ (GMS, AA 4:447). Kant’s claim is only that we must attribute freedom to the *concept* of a rational being with a will, and the question of whether we must attribute freedom to *ourselves* remains open.

Still, one might think we can reconstruct an argument for freedom on the basis of subsection two, regardless of how well it captures Kant’s intentions in *Groundwork* III. Indeed, a number of contemporary philosophers have thought there is a deeper truth to Kant’s assertion that a rational being must act ‘under the idea of freedom’ since this treats freedom as a ‘necessary condition of playing the game of deliberation’ (Hill 1998, 265) or as a ‘fundamental feature of the standpoint from which decisions are made’ (Korsgaard 1996b, 163). Along these lines, it seems Kant could have said:

- (1) We must play the game of deliberation.
- (2) In playing this game, we must think of ourselves under the idea of freedom.
- (3) A free will and a will under moral laws are co-entailing concepts.
- (4) Therefore, given (1)–(3), morality is valid for us.

Even as a reconstruction, it is not clear if this argument works since the sense of ‘freedom’ in premise (2) is too weak to support the inference to premise (3). To see why, suppose that I am caught between two desires, say, between working for the rest of the afternoon and going to the movies. We can agree that there is a sense of ‘freedom’ relevant to my decision-making: when I give reasons for my action, whether to work or go to the movies, I am guiding myself by concepts, and by doing so, I am presupposing my distance from the two desires I feel at the time. Accepting this, we are still not committed to saying that I must regard myself as independent from *all* external influences when I play the game of deliberation, for I could very well be calculating the comparative degree of pleasure I will receive from one activity versus the other. Nor are we committed to saying that I must regard myself as autonomous in this game, i.e. as legislating



myself by laws of reason, for the principles I appeal to in making up my mind may be prudential and so dependent on my natural sensibility after all. In short, the status of 'freedom' as a practical assumption – even if a true characterization of agency in general – falls short of both the negative and positive senses of freedom in premise (3), and so falls short of supporting the conclusion in (4).

### 5.2.2. *Second passage: the spontaneity of theoretical reason*

Aside from whether the above reconstruction works, those interested in capturing Kant's intentions in *Groundwork* III will no doubt press me on another passage where he appears to offer non-moral evidence for our freedom.<sup>36</sup> Consider what he says, for example, about the spontaneity of theoretical reason:

Now, a human being actually does find in himself a capacity by which he is distinguished from all other things, even from himself, in so far as he is affected by objects, and that is *reason* ... [This faculty] under the name of the ideas shows a spontaneity so pure that thereby he goes far beyond anything that sensibility can ever afford him. (GMS, AA 4:452; cf., KrV A546/B574; Prol AA, 4:345)

A few sentences earlier, Kant had claimed that 'what reaches our consciousness not by affection of the senses, but immediately' exhibits a 'pure activity' (*reine Thätigkeit*) within us, and that such activity indicates that we do not belong wholly to the world of sense. After identifying the source of this activity in our capacity to think ideas of reason, Kant then concludes: 'On account of this a rational being must view itself, as an intelligence (thus not from the side of its lower powers), as belonging not to the world of sense, but to that of understanding' (GMS, AA 4:452). Here, the argument seems to be something like this:

- (1) We are conscious of reason's pure self-activity within ourselves.
- (2) This activity secures our membership in the intelligible world.
- (3) As members of this world, we are negatively free (i.e. free from all external influences).
- (4) The principle of a negatively free will is the principle of self-legislation.
- (5) The principle of self-legislation is equivalent to morality.
- (6) Therefore, given (1)–(5), morality is valid for us.

Now it is unclear to me if this conclusion goes through since the spontaneity we display in thinking ideas is not yet a causality, that is, it is not yet freedom to *act* on pure practical reasons.<sup>37</sup> However, for our purposes here, I need not deny that Kant appeals to the spontaneity of theoretical reason in *Groundwork* III, and so I need not deny the textual evidence supporting premises (1)–(3). The question we must ask is whether these premises are meant to warrant our claim to freedom, as many commentators think, or whether they are merely intended to prepare the way for such a warrant. My own view is that the second option is more charitable and also more consistent with the range of texts we have examined above,<sup>38</sup> and while I lack space to explore this issue at length, it is certainly one that deserves further discussion.

### 5.3. Third objection: a non-moral route pre-1785

A final objection I would like to consider is that my reading lies in tension with the trajectory of Kant's pre-1785 writings, all of which suggest he was seeking a non-moral route to freedom. In the larger context of these writings, it seems that Kant hoped to derive freedom from theoretical reason, either by appealing to the spontaneity of 'I think' (Section 5.3.1) or by appealing to the spontaneity of the understanding (Section 5.3.2).

#### 5.3.1. First passage: the spontaneity of 'I think'

The first piece of evidence is from Kant's lectures on metaphysics during the mid-1770s:

When I say, 'I think, I act, etc.' either the word 'I' is employed improperly, or I am free [ ... ] But now I am conscious to myself that I can say: 'I do'; therefore I am conscious of no determination in me, and thus I act *absolutely freely* [ ... ] All practical objective propositions would have no sense if human beings were not free. (V-Met/Heinze, AA 28:269)

The present argument seems to be (i) that by expressing 'I think, I do' we display a pure form of spontaneity, (ii) that this spontaneity proves our freedom from all external influences, and (iii) that we therefore have a secure foundation for 'practical objective propositions,' i.e. for moral laws. As Dieter Henrich interprets this passage in his 1960 essay, Kant is proceeding from a strictly theoretical starting point. 'That moral obligation is something real,' Henrich explains, 'is certain because the consciousness of the self in thought, which alone could doubt that obligation, can be conceived only as freedom. The principle of theoretical reason gives foundation to the possibility of moral existence.'<sup>39</sup>

Other passages from this lecture show that Henrich's reading is less decisive than it first appears. At one point, Kant says freedom is beyond our comprehension, but that we still have grounds to attribute freedom to ourselves because we are subject to practical imperatives (V-Met/Heinze, AA 28:270). After claiming that 'all practical objective propositions would have no sense if human beings were not free,' he adds: 'But now there are such imperatives according to which I should do something; therefore all practical imperatives, problematic as well as pragmatic and moral, must presuppose a freedom in me' (V-Met/Heinze, AA 28:269). This point is all the more important when we return to the passage Henrich cites, for we can see that Kant is using a normative premise – the consciousness of our will under imperatives – to warrant our claim to 'absolute spontaneity' (V-Met/Heinze, AA 28:268). What makes Kant's argument in the 1770s different from his mature position, then, is not that he appeals to the spontaneity of thinking to infer our freedom in the practical sphere, as Henrich and others suppose. What makes his argument different is that he thinks we have access to our freedom via practical imperatives in general, not via moral imperatives in particular.

### 5.3.2. *Second passage: from thinking to willing*

Another place scholars have thought Kant was seeking a non-moral route to freedom is in his 'Review of Schulz' from 1783 – roughly the time he was planning to write the *Groundwork*. Near the end of this review, Kant says that even a strict fatalist

has assumed in the depths of his soul that understanding is able to determine his judgment in accordance with objective grounds that are always valid and is not subject to the mechanism of merely subjectively determining causes [ ... ] hence he always admits freedom to think, without which there is no reason. In the same way [*Eben so*] he must also assume freedom of the will in acting, without which there would be no morals. (RezSchulz, AA 8:14)

Scholars have read this passage as saying that the spontaneity we exhibit in the cognitive sphere (where we govern our understanding) warrants the assumption that we are free in the practical sphere (where we govern our will). However, a few sentences above the cited passage, Kant writes that the 'most confirmed fatalist, who is a fatalist as long as he gives himself up to mere speculation, must still, *as soon as he has to do with wisdom and duty*, always act as if he were free, and this idea also actually produces the deed that accords with it and can alone produce it' (RezSchulz, AA 8:13; my emphasis). Thus, Kant's point is that a fatalist must assume he is free as soon as he becomes conscious of the demands of morality (i.e. 'wisdom and duty'), for such demands presuppose his independence from sensible inclinations. Compared to his lectures on metaphysics, it is clear that by 1783, Kant had attained a version of his mature position – that only moral laws provide grounds for cognizing our freedom – since he must have realized that technical and prudential imperatives do not presuppose our absolute spontaneity.

If anything, the passages I have discussed above show that Kant's commitment to finding moral evidence for freedom is much earlier than commentators have assumed. In addition to the 'Review of Schulz' from 1783,<sup>40</sup> we find Kant arguing along these lines in the 1781 edition of the first *Critique*:

I assume that there are really pure moral laws, which determine completely a priori (without regard to empirical motives, i.e. happiness) the action and omission, i.e. the use of the freedom of a rational being in general, and that these laws command absolutely (not merely hypothetically under the presupposition of other empirical ends), and are thus necessary in every respect. *I can legitimately presuppose this proposition by appealing not only to the proofs of the most enlightened moralists but also to the moral judgment of every human being, if he will distinctly think such a law.* (KrV A807/B835; my emphasis)<sup>41</sup>

By 1781, then, we have evidence that Kant believed moral laws offer us grounds for cognizing our freedom, and what is more, we have evidence that he considered the 'judgment of every human being' sufficient to assume the existence of such laws.<sup>42</sup> Of course, it is difficult to say precisely when Kant adopted this view since the lectures on metaphysics from the mid-1770s suggest he had yet to see the special status of moral imperatives compared to technical and prudential

imperatives. From the documents we have of this time, signs foreshadowing Kant's mature view begin to appear in the late 1770s, and on inspection, it is striking how close they come to stating a thesis normally associated with the doctrine of the Fact of Reason. In *Reflexion* 6849, for example, Kant writes that the '**primary ought** (original = *absolute* or the universal idea of duty) cannot be comprehended,' adding in *Reflexion* 6850: 'The primary ought is a condition under which alone freedom becomes a capacity in accordance with constant rules that determine a priori.' But even if we treat such literary fragments as inconclusive, the fact remains that Kant's moral orientation shines through in the early 1780s. And we should not overlook the fact that in the *Groundwork* itself, Kant states clearly that reason '*first becomes aware* that it can of itself also be practical' through 'the pure thought of *duty* and in general of the *moral law*, mixed with no foreign addition of empirical inducements' (GMS, AA 4:410; my emphasis). So whether we focus on the official or unofficial documents of this time, there is ample evidence to believe Kant was committed to a moral argument for freedom much earlier than 1788.

## 6. The real difference between Kant's two works

If the interpretation I have defended above is correct, then we have no reason to believe Kant changed his approach to moral justification between 1785 and 1788. Instead, we have good grounds to take Kant at his word, as readers did for many decades, when he claimed that the second *Critique* 'presupposes' the *Groundwork* in its formulation and justification of the moral principle. As we have seen, the real difference between Kant's two works is a matter of presentation: the second *Critique* says little about the formula of the moral law, and even less about how it appears to us as an 'ought.' In my view, we can explain this difference in terms of the two genres Kant self-consciously worked in: classical moral philosophy, on the one hand, and critical transcendental philosophy, on the other hand.

In the Preface to the second *Critique*, Kant tells us his goal is to show '*that there is pure practical reason*' (KpV, AA 5:3). This differs from his stated goal in the *Groundwork*, namely, to explain and establish '*the supreme principle of morality*' (GMS, AA 4:392). Now if the second *Critique* presupposes the *Groundwork*, we should not be surprised to discover abbreviated versions of the 1785 work turning up in the 1788 work. Bearing in mind Kant's different aims, we can see why he would devote an entire section to explaining the principle of morality as a law of autonomy in *Groundwork* II, only to compress this discussion in the second *Critique*. In the latter work, Kant bypasses his complex analysis of the formulas, and proceeds directly to what he calls 'the fundamental law' of pure practical reason: 'So act that the maxim of your will could always hold at the same time as a principle in a giving of universal law' (KpV, AA 5:30). Similarly, we can also see why Kant would devote an entire subsection to establishing

the principle of morality as a categorical imperative in *Groundwork* III, only to compress this discussion in the second *Critique*. In the latter work, Kant bypasses the details of his second deduction, and proceeds directly to say that morality takes the form of an imperative for us because we are, unlike *holy* wills, imperfectly rational (KpV, AA 5:32).

My hypothesis, then, is that the *Groundwork* and the second *Critique* diverge in their presentations because they belong to different genres of exposition. On the one hand, the *Groundwork* was written to prepare for a ‘future metaphysics of morals,’ a project Kant had conceived for many years but only got around to publishing in 1797. What this shows is that Kant’s *Groundwork* belongs to the tradition of classical moral philosophy: it aims to support a system of normative principles that could inform real-life decision-making.<sup>43</sup> What Kant must have realized was necessary before the publication of such a system was a clearing of foundations – in particular, a *clearing away* of popular moral philosophy – in order to articulate the true method of moral inquiry (cf., KrV Axx). On the other hand, the second *Critique* was written to complete a project Kant had already begun in 1781 with his first *Critique*. What this shows, by way of comparison, is that the second *Critique* belongs to the genre of Kant’s own critical philosophy: it aims to show that pure reason, though wholly dialectical in the speculative sphere, has a legitimate use in the practical sphere. From the perspective of genres, Kant’s brevity in the second *Critique* makes perfect sense. He did not feel compelled to offer a long analysis of the categorical imperative, nor a deduction of its possibility, because his aim was no longer to prepare for a future metaphysics of morals. His aim was to render the transition from pure speculative to pure practical reason more perspicuous, and to that end, his first objective was simply to show *that* pure practical reason exists.<sup>44</sup>

## 7. Closing remarks

What I hope to have achieved in this paper is a plausible alternative to the standard story of Kant’s deductions of morality and freedom. After showing that it is relatively new in the reception of Kant’s ethics (Section 2), I suggested that contemporary readers have found this story compelling because it purports to explain Kant’s worry about a ‘hidden circle’ in *Groundwork* III (Section 3.1). On my reading, the idea of an ‘intelligible world’ avoids the circle, but only because our consciousness of its laws, i.e. the laws of morality, legitimates our claim to freedom (Section 3.2). Turning to the second *Critique*, I showed that Kant’s references to the *Groundwork* are positive on every score, and that his comment about a ‘vainly sought deduction’ of the moral law is consistent with the conclusion of *Groundwork* III (Section 4.1). With these details in place, I offered my full defense of the Continuity Reading, showing that the arguments of *Groundwork* III and the second *Critique* share the same general structure (Section 4.2). Then, after

defending my reading from three potential objections (Section 5), I offered a hypothesis for why Kant's two works still differ in their presentation (Section 6).

Questions of interpretation aside, it is difficult to say how we should evaluate Kant's project of justification today, if not in terms of his *failure* to derive the moral law from non-moral premises. This impression is only reinforced by the fact that a number of contemporary Kantians have thought we can reconstruct a successful defense of the moral law from an account of action and agency in general.<sup>45</sup> In this light, Kant's admission in the second *Critique* that no deduction of the moral law is possible, even if consistent with his earlier work, appears to fall short of what we ultimately want: a strict philosophical proof. At the same time, we must not lose sight of the fact that Kant drew attention to the limits of his argument in *Groundwork* III, saying that it is 'a reproach that must be brought against human reason in general, that it cannot make comprehensible as regards its absolute necessity an unconditional practical law (such as the categorical imperative must be)' (GMS, AA 4:463). Whether or not the reproach is just, I do not think we should assume in advance that our ideal strategy is one of finding a non-moral route to freedom. Kant believed our consciousness of the moral law is the only viable orientation a project of justification can take because it is the only orientation that works *within* the bounds of human reason.

At this final juncture, it may seem that we are left with a standoff of intuitions about what the task of justification in ethics should accomplish. Those who think it should refute all forms of moral skepticism will likely find my version of the Continuity Reading disappointing. If I am right, Kant was not seeking to justify the necessity of the moral law; indeed, that is precisely what he thought goes beyond our comprehension and so beyond the scope of any deduction. Those who insist that we should be seeking such a deduction will wonder what kind of 'justification' of morality Kant ends up with in the end, if any. To call our awareness of the moral law a 'fact' hardly seems satisfying, even if we follow Kant's repeated claim that this fact is part of our common moral consciousness. However, it is worth emphasizing that Kant did not regard refuting the skeptic to be the measure of success for his moral philosophy. He was very much speaking to those of us who are committed to morality, but who lack the self-understanding necessary to make this commitment intelligible. That in itself does not settle the standoff of intuitions we have, but it does bring an important point to light. It shows that how we frame the task of justification in ethics comes down to where we locate the center of our discussion – whether in philosophical speculation or in ordinary life. In arguing *from* morality, Kant was, I believe, attempting to begin with the latter.



## Abbreviations

In the case of the *Critique of Pure Reason*, I follow the standard practice of referring to the 1781 (A) and 1787 (B) editions. For all other texts, citations appear in the order of abbreviation, volume number, and page number from the *Akademie Ausgabe* (AA), *Kants Gesammelte Schriften*, edited by *Königlich Preussische Akademie der Wissenschaften* (29 vols. Berlin: de Gruyter, 1900–). Translation decisions are my own, though I have consulted (and sometimes followed) *The Cambridge Edition of the Works of Immanuel Kant*, edited by Paul Guyer and Allen Wood (Cambridge: Cambridge University Press, 1992–).

Br	<i>Correspondence</i>
FM	<i>What Progress has Metaphysics Made in Germany?</i> (1791)
GMS	<i>Groundwork for the Metaphysics of Morals</i> (1785)
KpV	<i>Critique of Practical Reason</i> (1788)
KrV	<i>Critique of Pure Reason</i> (1781/1787)
KU	<i>Critique of Judgment</i> (1790)
MS	<i>Metaphysics of Morals</i> (1797)
ProI	<i>Prolegomena to Any Future Metaphysics</i> (1784)
RGV	<i>Religion Within the Boundaries of Mere Reason</i> (1793)
Refl	<i>Fragments and Reflections</i>
RezSchulz	'Review of Schulz' (1783)
VAMS	'Preparatory Work for the <i>Metaphysics of Morals</i> '
V-Met/Heinze	<i>Lectures on Metaphysics: Heinze</i>
V-MS/Vigil	<i>Lectures on Ethics: Vigilantius</i>
VT	'On a Newly Raised Superior Tone in Philosophy' (1796)

## Notes

1. For the standard view, see Ross (1954); Beck (1960); Henrich (1960); Williams (1968); Ameriks (1981); Korsgaard (1989); O'Neill (1989); Allison (1990); Łuków (1993); Neiman (1994); Hill (1998); Allison (2011); Rawls (2000); Engstrom (2002); Darwall (2006); Sussman (2008); Reath (2012); and Grenberg (2013). Other scholars find a significant change in Kant's strategy of justification, not because he gave up a proof of the moral law, but because he gave up a non-moral argument for freedom. For this view, see Schönecker (1999; 2006; 2013; 2014); Guyer (2009); Ludwig (2010); Timmermann (2010); Ludwig (2010; 2012; 2015); Hahmann (2012); Bojanowski (forthcoming).
2. My claim that Kant's project of justification in ethics is continuous between *Groundwork* III and the second *Critique* is not new. Some have defended a strong version of this reading, according to which Kant's two works are identical (McCarthy 1982); others have defended a more moderate version, according to which Kant foreshadows, without developing, the doctrine of the 'Fact of Reason' in 1785 (Paton 1947; Henrich 1975). What makes my position in this paper novel is that, while I am willing to concede that Kant's two works still differ (contrary to the strong version), I shall argue that a moral argument for freedom is already present in *Groundwork* III (contrary to the moderate version).

3. That Kant remained committed to this argument *after* 1787 is clear from various published and unpublished texts. For a few examples, see Refl 7316, 7321; KU, AA 5:275; RGV, AA 6:26n, 6:49n; FM, AA 20:345; VT, AA 8:403; V-MS/Vigil, AA 27:506; VAMS, AA 23: 245; and MS, AA 6:252.
4. To be clear, the argument from freedom to morality says that we have grounds independent of the moral law to think we are free, and those grounds explain the possibility of the moral law as the supreme principle of our will. By contrast, the argument from morality to freedom says that we have no insight into the possibility of the moral law as the supreme principle of our will, but that our consciousness of this law nevertheless grounds our claim to freedom. One obvious criticism facing the latter argument, and often leveled against Kant's doctrine of the Fact of Reason, is that it slips into dogmatism about the moral law's normative authority. Due to limits of space, however, I will not be addressing this issue here. I discuss it at greater length in Ware (2014).
5. Other advocates of this view include Wolff (2009), who defends continuity between Kant's works by comparing Kant's strategy of moral justification to his theory of geometrical postulates (2009), and Puls (2011; 2014), who defends continuity by highlighting the role of common human reason in *Groundwork* III. While I am in broad agreement with their accounts, I shall be seeking a new line of interpretation in this paper.
6. In his *Letters on Kantian Philosophy* from 1786/87, the term '*Tatsache*' occurs numerous times (100, 109, 125, 126). For an excellent study of Reinhold's role in the development of post-Kantian philosophy, see Ameriks (2000).
7. See his (1796–97, vol. I, 139). For another example, see Brastberger's (1792, 57, 71).
8. See his (1796–97, vol. I, 207, 247n, 255, 282, and 352).
9. See Tittel (1786); Pistorius (1786a; 1786b); and Stattler (1788).
10. See his (1800, 174–175).
11. This quotation comes from his (1844, 535).
12. This essay was translated in 1994 under the title 'The Concept of Moral Insight and Kant's Doctrine of the Fact of Reason.'
13. Herbert Paton reached a similar conclusion in his book, *The Categorical Imperative: A Study in Kant's Moral Philosophy*, published in 1947. Curiously, Paton expressed a great degree of ambivalence on the topic of whether Kant changed his proof-structure in the second *Critique*. He believed Kant was seeking a strict proof in *Groundwork* III that would derive the validity of the moral law from an 'independent and necessary presupposition of freedom,' a deduction he subsequently gave up in 1788. But in the very next paragraph Paton said, we need not 'exaggerate the difference between the two views' (1947, 203), adding: 'it is doubtful how far he regards freedom in the *Groundwork* as a non-moral principle' (204). What is even more odd is that Paton cited the final paragraph of *Groundwork* III where Kant states the impossibility of deducing the moral law from theoretical premises (GMS, AA 4:463). Yet, Paton took this as evidence that Kant was merely 'beginning to see dimly' that his proof was doomed to fail (225).
14. See his (1994, 81). Henrich appears to have changed his mind in a later essay when he claimed that Kant's argument in *Groundwork* III has an irreducibly moral core (1975, 86). See also Ameriks's recent comments (2012, 24–25) on his earlier (1981) defense of the Reversal Reading.
15. Barbara Herman's observation that in the 'resurgence of work on Kant's ethics, one notices the quiet avoidance of the issue of justification' sadly rings true today (1989, 131), especially in the English scholarship. Only recently has the German scholarship shown a resurgence of work on *Groundwork* III and the

second *Critique*. See, for example, the two collections edited by Puls (2014) and Schönecker (2015).

16. The status of the circle in *Groundwork* III is worthy of its own treatment, and here I can only touch the surface of a much larger issue. For further discussion, see Schönecker (1999); Allison (2011); and Berger (2015).
17. With the exception of Wolff (2009, 546, note 64) and Puls (2011, 543, note 19), the significance of this remark has been neglected in the literature devoted to *Groundwork* III.
18. I offer a more detailed account of the transition-structure of the *Groundwork* in Ware (2016).
19. We find a nearly identical set of claims in the second *Critique*, but there Kant speaks of an 'archetypal world,' which he says is a 'model [*Muster*] for the determinations of our will' (KpV, AA 5:43).
20. One might be open to this claim but still wonder why such an ideal must be a recognizably *moral* one (as is the 'Fact of Reason') as opposed to a more generic ideal of rationality or agency. As an anonymous reviewer puts it: 'Kant's argument in *Groundwork* III might just be that once we appreciate the connections between morality, autonomy, and rationality, we can also see that moral acts are the only possible fully rational or intelligible or justified ones, or those in which our agency is most fully realized.' The reason why I wish to insist upon calling the concept of an intelligible world a moral ideal is that it specifies practical laws that, in contrast to technical or prudential imperatives, express absolute necessity. Moreover, Kant himself is clear in the first *Critique* that the intelligible world is a moral ideal, so my reading has good textual support. As he states, the intelligible world is nothing more than the idea of the 'world as it would be if it were in conformity with all moral laws' (KrV A808/B836; see also Refl 5086 and 6977 dated from the 1770s). For an illuminating account of this passage in connection with *Groundwork* III, see Rauscher (2009; 2015).
21. It is worth pointing out that when Kant introduces the concept of an intelligible world in *Groundwork* III, he says that a human being has 'two standpoints from which he can regard himself and cognize [*erkennen*] laws for the use of his powers,' by which he means to include our practical powers of will and volition (GMS, AA 4:452). As members of the world of sense, Kant explains, we have a standpoint to consider ourselves 'under laws of nature,' but as members of the intelligible world, we have a standpoint to consider ourselves 'under laws which, being independent of nature, are not empirical *but grounded merely in reason*' (GMS, AA 4:452; my emphasis). Since laws 'grounded merely in reason' are none other than the laws of morality (cf., GMS, AA 4:458 and 4:462), the implication of this passage is very similar to what Kant will assert in the second *Critique*: that our moral consciousness reveals our freedom of will.
22. We find a similar claim in the second *Critique*: 'Yet we are conscious through reason of a law to which all our maxims are subject, as if a natural order must at the same time arise from our will. This law must therefore be the idea of a nature not given empirically and yet possible through freedom, hence a supersensible nature to which we give objective reality at least in a practical respect, since we regard it as an object of our will as pure rational beings' (KpV, AA 5:44; cf., KrV B430-31).
23. At this 'outermost boundary' of practical inquiry, what remains beyond our epistemic reach is not the moral law's conceptual structure. Kant had already clarified that structure in *Groundwork* II using a method of analysis. What remains 'incomprehensible' then is the moral law's normative authority. Interestingly,

Kant repeats this point in the second *Critique* when he explains why the moral law, unlike the pure categories of the understanding, admits of no justification: 'With the *deduction* [of the moral law], that is, the justification of its objective and universal validity and the discernment of the possibility of such a synthetic proposition a priori, one cannot hope to get on so well as was the case with the principles of the pure theoretical understanding' (KpV, AA 5:46). See Baum (2014) for further discussion.

24. My understanding of Kant's second deduction has been influenced by Timmermann (2007) and Stern (2012; 2015), who both argue – correctly, I believe – that in this section of *Groundwork* III, Kant is only trying to explain the possibility of moral bindingness (i.e. the imperative form the moral law takes to finite rational beings like us), not the moral law itself. For a helpful discussion of this distinction, i.e. between the moral law and its imperative form, see Willaschek (1992, 176).
25. Many commentators agree that Kant's appeal to the idea of an intelligible world plays a crucial role in *Groundwork* III, but what they overlook is that this idea has two functions. When we consider ourselves as members of an intelligible world, we have grounds to ascribe a *negative* sense of freedom to ourselves since we are presupposing our independence from the world to which our impulses and inclinations belong, the world of sense. Yet, on my reading, this is not yet the key step. The key step is that the intelligible world also affords us a standpoint for thinking of ourselves in accordance with a pure practical law, i.e. the moral law. Only now do we have grounds to ascribe a *positive* sense of freedom to ourselves – to think of ourselves as possessing a will – since the moral law is self-legislated. Note, too, that at GMS, AA 4:458 Kant says that these two senses of freedom are 'bound' (*verbunden*) in the idea of an intelligible world, reiterating his earlier stipulation that synthetic propositions are 'only possible by this, that both cognitions are bound [*verbunden*] together by their connection with a third, in which they are both to be found' (GMS, AA 4:447; cf., KrV B162-65).
26. Timmermann, for example, has argued that Kant's mention of a 'vainly sought deduction' (*vergeblich gesuchten Deduction*) is a thinly veiled reference to *Groundwork* III: 'Of course, Kant does not explicitly say that *he* is the person who tried in vain to provide a deduction of the moral principle, but as *Groundwork* III contains precisely such a deduction, and proving the possibility of the categorical imperative was hardly fashionable amongst his philosophical colleagues, we can safely infer that Kant had come to reject his earlier justificatory attempt by the time he composed the second *Critique*' (2010, 74). On closer inspection, however, this assumption is far from self-evident. Just before Kant says that 'something different and quite paradoxical takes the place of this [*dieser*] vainly sought deduction,' he writes: '[T]he objective reality of the moral law cannot be proved by *any* deduction, by *any* efforts of theoretical reason, *speculative* or *empirically supported*' (KpV, AA 5:47; my emphasis). This suggests Kant's reference is quite general and does not necessarily imply his earlier argument. Given its placement in the text, the 'vainly sought deduction' is only a reproach against efforts to establish the validity of the moral law on theoretical grounds, i.e. the sort we find in the writings of Hobbes, Hutcheson, Wolff, or Crusius.
27. Kant also echoes various turns of phrase from the last paragraph of *Groundwork* III when he writes: 'But instead of the deduction of the supreme principle [*obersten Princip*s] of pure practical reason – that is, the explanation of the possibility of such a cognition a priori – nothing more could be adduced than that, if one had insight into the possibility of freedom of an efficient cause, one would also have

insight into not merely the possibility but even the necessity of the moral law [ ... ] But no insight can be had into the possibility of the freedom of an efficient cause, especially in the sensible world: we are fortunate [*glücklich*] if only we can be sufficiently assured that there is no proof of its impossibility, and are now forced [*genöthigt*] to assume it and are thereby justified in doing so by the moral law, which postulates it' (KpV, AA 5:93–94).

28. For an insightful discussion of these points, see Wood (2011). Schönecker defends a different reading, arguing that Kant is only speaking of justifying a 'determinate formula' of the moral principle, the project of *Groundwork* II (2013, 8, note 10). Yet, on examination, Kant's remark proves to be ambiguous. When he speaks of justifying a 'determinate formula' of the moral principle, he could mean either (a) justifying the conceptual structure of the principle, or (b) justifying the principle's application to us as a categorical imperative. I think (b) is a plausible reading, especially since Kant never characterizes the task of *Groundwork* I or II in terms of '*Rechtfertigung*' (and elsewhere, when Kant invokes this particular term, he often addresses the synthetic application of a concept, not its mere analytic explication). That being said, even if Schönecker is right, Kant's exclusion of *Groundwork* III from what the second *Critique* 'presupposes' does not offer direct support for the view that he abandoned his earlier strategy of justification. So even if we read Kant's remark in terms of (a), further evidence is required to support the Reversal Reading.
29. Aside from the passage at KpV, AA 5:8, there is only one other place in the second *Critique* where Kant alludes to his earlier strategy of justification. As rational beings, he writes, we can regard ourselves as members of an intelligible order, 'for it has been *sufficiently proved elsewhere* that freedom, if it is attributed to us, transfers us into an intelligible order of things' (KpV, AA 5:42; my emphasis). Here, Kant is referring to Step 1 of his deduction from *Groundwork* III, that if we think of ourselves as free from external influences, we must take up a 'different standpoint' than when we 'represent ourselves in terms of our actions as effects that we see before our eyes' (GMS, AA 4:451). But we must ask: Why would Kant cite this step with approval in the second *Critique*? If he had reversed his structure of justification, consistency would have forced him to reject this earlier means of escaping the circle between freedom, autonomy, and morality. Instead of rejecting this move, though, we find Kant speaking of the doctrine of the Fact of Reason and his earlier strategy in the very same paragraph, as if the two were entirely compatible (KpV, AA 5:42).
30. Kant's motive for making the doctrine of the Fact of Reason prominent in the second *Critique* was perhaps a reaction to Tittel's (1786). Kant replied to Tittel by saying: 'A reviewer who wanted to say something censuring this work [i.e. the *Groundwork*] hit the mark better than he himself may have intended when he said that no new principle of morality is set forth in it but only a new formula. But who would even want to introduce a new principle of all morality and, as it were, first invent it? Just as if, before him, the world had been ignorant of what duty is or in thoroughgoing error about it' (KpV, AA 5:8n). By stressing that our consciousness of the moral law is something actual – a 'fact' that precedes abstract speculation – Kant likely wanted to prevent other readers from confusing the formula of morality with morality itself, as Tittel did. It is also telling that the footnote at KpV, AA 5:8 appears just after Kant speaks of his earlier attempt to *justify* the formula of morality in *Groundwork* III.
31. Kant continues: 'Accordingly the moral law is for them an *imperative* that commands categorically because the law is unconditional; the relation of such



a will to this law is *dependence* under the name of obligation, which signifies a *necessitation* [Nötigung], though only by reason and its objective law, to an action which is called *duty*' (KpV, AA 5:32).

32. As an anonymous reviewer points out, one might also suspect that the different titles Kant uses in *Groundwork* III and the second *Critique* reflect a deeper shift in his project. In the former work, he speaks of a 'Critique of pure practical reason,' only to claim in the latter work that we have 'no need to criticize the *pure faculty itself* in order to see whether reason is merely making a claim in which it presumptuously *oversteps* itself' (KpV, AA 5:3). Now this would indeed signal a major shift if Kant believed in *Groundwork* III that pure practical reason oversteps its boundaries (and so requires a 'Critique' in the negative sense of the word). However, when we turn to the details of the text, we find Kant speaking of a 'Critique' in the positive sense of the word, i.e. an examination of reason's sources and limits (c.f., KrV A11/B25). These two senses of 'Critique' are also implicit in Kant's account of why he introduced the ideal of an intelligible world in *Groundwork* III. That ideal, he writes, served merely to '*limit* the principle of motives from the field of sensibility by circumscribing this field and showing that it does not include everything within itself but that there is still *more* beyond it' (GMS, AA 4:462; my emphasis). In other words, by curbing the pretensions of empirical practical reason (a negative Critique), Kant is able to show that practical reason has a pure use as well (a positive Critique). All of this coheres with his reasons for calling the second *Critique* a 'Critique of practical reason.' For a different reading, see Bojanowski (forthcoming).
33. I thank an anonymous reviewer for formulating this objection.
34. This shifts the burden of proof onto the standard reading. For if we assume that the 'demonstrable proposition' refers to the moral law's supreme normative authority, then we are forced to say that Kant contradicts himself when (at GMS, AA 4:463) he claims that such authority lies beyond our epistemic reach.
35. For further discussion of this passage, see Tenenbaum (2012); Saunders (2014); and Horn (2015).
36. Ameriks (2003) and Guyer (2009), for example, have pursued a reading of *Groundwork* III along these lines.
37. Another potential problem with this argument – noted first by Pistorius in his 1786 review of the *Groundwork* – is that Kant appears to overstep the limits he places on human knowledge by resting his deduction of the categorical imperative on a noumenal claim: '[Kant] helps himself to his problematic concept of freedom, transfers us from the world of sense to the world of understanding and brings over from this completely unknown world [völlig unbekannten Welt] the grounds of the possibility and necessity of his categorical imperative' (Pistorius 1786a, 159; my emphasis; see also Pistorius 1786b, 16). For a helpful discussion of Pistorius's charge and its potential impact on Kant, see Ludwig (2010; 2012). See also Schönecker (1999, 2006) and Guyer (2007) for recent reformulations of this line of criticism. While this topic deserves fuller treatment, I would here like to make one preliminary remark. Even if we agree that Kant was sensitive to Pistorius's review, it remains an open question whether his reaction was (i) to rewrite the inner structure of his argument in 1788, or (ii) to change its point of focus. The fact that Kant was influenced by Pistorius does not by itself resolve this issue.
38. In yet another important passage, Kant writes that the idea of freedom 'is valid only as a necessary presupposition of reason in a being that believes itself to be conscious of a will, i.e. of a capacity distinct from a mere faculty of desire



(namely, to determine itself to action as an intelligence, hence according to laws of reason independently of natural instincts' (GMS, AA 4:459). What is noteworthy here is that Kant does not say the idea of freedom is valid for a being conscious of a capacity to think pure theoretical ideas. Nor does he say the idea is valid for a being conscious of a capacity to act according to prudential imperatives. Instead, Kant is careful to qualify his remark so that the idea of freedom is valid only for a being conscious of a capacity to act according to laws of reason, that is, according to moral laws.

39. Cited from the English translation (1994, 80–81).
40. 1783 is the same year Kant published his *Prolegomena*, and in this work he draws an explicit link between freedom and the concept of an 'ought' (Prol, AA 4:344–245).
41. Some passages from the 1781 edition of the first *Critique* indicate that Kant had yet to appreciate the distinctive status of moral imperatives. At times, for instance, he talks as if our faculty of reason displays absolute spontaneity in all 'ought' statements, including prudential ones: 'Whether it is an object of mere sensibility (the agreeable) or even of pure reason (the good), reason does not give in to those grounds which are empirically given, and it does not follow the order of things as they are presented in intuition, but with complete spontaneity it makes its own order according to ideas' (KrV A547). However, following this passage, Kant goes on to say that by virtue of its ideas, reason 'declares actions to be necessary that yet have not occurred and perhaps will not occur, nevertheless presupposing of all such actions that reason could have causality in relation to them.' And later, he clarifies that the projected 'order' reason fashions for itself is that of a *moral* world, that is, a world of rational beings 'insofar as their free choice under moral laws has thoroughgoing systematic unity in itself as well as with the freedom of everyone else' (KrV A808/B836). In retrospect, Kant may have thought it was sufficient to point to practical imperatives in general to indicate a divide between what 'is' and what 'ought' to be, all the while knowing that only moral imperatives presuppose our complete spontaneity in the strict (i.e. transcendental) sense of the word. Thanks to Bernd Ludwig for pressing me on this issue.
42. In addition to Puls (2014), the centrality of 'common' moral self-understanding is discussed in detail by Grenberg (2013), Ware (2014), and Sticker (2015) in connection with Kant's Fact of Reason.
43. As Kant states in the Preface of the *Groundwork*: 'my aim here is actually directed to moral world-wisdom [*sittliche Weltweisheit*]' (GMS, AA 4:389).
44. The genesis of the second *Critique* bears out my hypothesis. In June 1787, Kant reported to be working on a new monograph titled a *Critique of Practical Reason*. In a letter to Friedrich Gottlob Born, he wrote: 'This work will better demonstrate and make comprehensible the possibility of supplementing, by pure practical reason, that which I denied to speculative reason' (Br, AA 10:490). Similarly, in a letter to Reinhold, dated at the end of December 1787, he wrote: 'This little book will sufficiently resolve the many contradictions that the followers of the old-guard philosophy imagine they see in my Critique [of Pure Reason]' (Br, AA 10:513–16). For a superb discussion of this history, see Klemme (2010).
45. I am thinking of Nagel (1970), Velleman (1989; 2009), and especially Korsgaard (1996a; 2009).

## Acknowledgements

For helpful feedback on previous incarnations of this paper, I wish to express my gratitude to Gabriele Gava, Andree Hahmann, Dai Heide, Bernd Ludwig, Sven Nyholm, Joe Saunders, Martin Sticker, Sergio Tenenbaum, Evan Tiffany, Leah Ware, Marcus Willaschek, Benjamin Yost, Ariel Zylberman, two anonymous *CJP* reviewers, as well as audience members at the Goethe-Universität Frankfurt and the Georg-August-Universität Göttingen. Thanks also to the Alexander von Humboldt-Stiftung for offering me a fellowship during the final stages of writing this paper.

## Funding

This work was supported by the Alexander von Humboldt-Stiftung.

## Notes on Contributor

**Owen Ware** is an Assistant Professor in Philosophy at Simon Fraser University. He works primarily on Kant, 19<sup>th</sup>-century German philosophy, and ethics.

## References

- Allison, Henry E. 1990. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Allison, Henry E. 2011. *Kant's Groundwork for the Metaphysics of Morals: A Commentary*. Oxford: Oxford University Press.
- Ameriks, Karl. 1981. "Kant's Deduction of Freedom and Morality." *Journal of the History of Philosophy* 19: 53–79.
- Ameriks, Karl. 2000. *Kant and the Fate of Autonomy*. Cambridge: Cambridge University Press.
- Ameriks, Karl. 2003. *Interpreting Kant's Critiques*. Oxford: Oxford University Press.
- Ameriks, Karl. 2012. *Kant's Elliptical Path*. Oxford: Oxford University Press.
- Baum, Manfred. 2014. "Sittengesetz und Freiheit. Kant 1785 und 1788." [Moral Law and Freedom: Kant 1785 and 1788.] In *Kants Rechtfertigung des Sittengesetzes in Grundlegung III: Deduktion Oder Faktum?*, edited by Heiko Puls, 209–226. Berlin: Walter de Gruyter.
- Beck, Lewis White. 1960. *A Commentary on Kant's Critique of Practical Reason*. Chicago, IL: University of Chicago Press.
- Berger, Larissa. 2015. "Der 'Zirkel' im dritten Abschnitt der Grundlegung – Eine neue Interpretation und ein Literaturbericht." [The Circle in the Third Section of the Groundwork – A New Interpretation and Literature Review.] In *Kants Begründung von Freiheit und Moral in 'Grundlegung III': neue Interpretationen*, edited by Dieter Schönecker, 9–82. Münster: Mentis.
- Bojanowski, Jochen. (Forthcoming). "Kant on the Justification of Moral Principles." *Kant-Studien*.
- Brastberger, Gebhard Ulrich. 1792. *Untersuchungen über Kants Kritik der praktischen Vernunft [Investigations on Kant's Critique of Practical Reason]*. Tübingen: Cotta'schen Buchhandlung.
- Darwall, Stephen. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.

- Engstrom, Stephen. 2002. "Introduction." In *Critique of Practical Reason*, edited and translated by Werner Pluhar, XV–ILV. Indianapolis, IN: Hackett.
- Fichte, Johann Gottlieb. 1794/95. *Grundlage der gesamten Wissenschaftslehre: als Handschrift für seine Zuhörer [Foundation of the Entire Science of Knowledge: As a Manuscript for His Listeners]*. Leipzig: Gabler.
- Grenberg, Jeanine. 2013. *Kant's Defence of Common Moral Experience: A Phenomenological Account*. Cambridge: Cambridge University Press.
- Guyer, Paul. 2007. *Kant's Groundwork for the Metaphysics of Morals: A Reader's Guide*. London: Routledge.
- Guyer, Paul. 2009. "Problems with Freedom: Kant's Argument in *Groundwork* III and its Subsequent Emendations." In *Kant's Groundwork of the Metaphysics of Morals: A Critical Guide*, edited by Jens Timmermann, 176–202. Cambridge: Cambridge University Press.
- Hahmann, Andree. 2012. "Ist 'Freiheit die Wahrheit der Notwendigkeit'? Das Ding an sich als Grund der Erscheinung bei Kant." [Is 'Freedom the Truth of a Necessity'? The Thing in Itself as a Ground for Appearance in Kant.] In *Sind wir Bürger zweier Welten? Freiheit und moralische Verantwortung im transzendentalen Idealismus*, edited by Mario Brandhorst, Andree Hahmann and Bernd Ludwig, 135–154. Hamburg: Meiner.
- Hegel, G. W. F. 1844. *Vorlesungen über die Philosophie der Geschichte [Lectures on the History of Philosophy]*. Berlin: Eduard Gans.
- Henrich, Dieter. 1960. "Der Begriff der sittlichen Einsicht und Kants Lehre vom Faktum der Vernunft." [The Concept of Moral Insight and Kant's Doctrine of the Fact of Reason.] In *Die Gegenwart der Griechen im neueren Denken. Festschrift für Hans-Georg Gadamer zum 60. Geburtstag*, edited by Dieter Henrich, W. Schulz and K.-H. Volkmann-Schluck, 130–172. Tübingen: Mohr Siebeck.
- Henrich, Dieter. 1975. "Die Deduktion des Sittengesetzes: Über die Gründe der Dunkelheit des letzten Abschnittes von Kants 'Grundlegung zur Metaphysik der Sitten.'" [The Deduction of the Moral Law: On the Reasons for the Obscurity of the Final Section of Kant's 'Groundwork for the Metaphysics of Morals.']. In *Denken im Schatten des Nihilismus. Festschrift für Wilhelm Weischedel zum 70. Geburtstag*, edited by Alexander Schwan, 55–112. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Henrich, Dieter. 1994. "The Concept of Moral Insight and Kant's Doctrine of the Fact of Reason." In *The Unity of Reason: Essays on Kant's Philosophy*, translated by Manfred Kuehn, edited by Richard L. Velkley, 55–87. Cambridge, MA: Harvard University Press.
- Herman, Barbara. 1989. "Justification and Objectivity: Comments on Rawls and Allison." In *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum*, edited by Eckart Förster, 131–144. Stanford, CA: Stanford University Press.
- Hill, Thomas E. 1998. "Kant's Argument for the Rationality of Moral Conduct." In *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*, edited by Paul Guyer, 249–272. New York, NY: Roman & Littlefield.
- Horn, Christoph. 2015. "Das Bewusstsein, unter dem moralischen Gesetz zu stehen – Kants Freiheitsargument in *GMS* III." [The Consciousness to Stand Under the Moral Law – Kant's Freedom-Argument in *GMS* III.] In *Kants Begründung von Freiheit und Moral in 'Grundlegung III': Neue Interpretationen*, edited by Dieter Schönecker, 137–156. Münster: Mentis.
- Klemme, Heiner. 2010. "The Origin and Aim of Kant's *Critique of Practical Reason*." In *Kant's 'Critique of Practical Reason': A Critical Guide*, edited by Andrews Reath and Jens Timmermann, 11–30. Cambridge: Cambridge University Press.
- Korsgaard, Christine. 1989. "Morality as Freedom." In *Kant's Practical Philosophy Reconsidered*, edited by Yirmiyahu Yovel, 23–48. Dordrecht: Kluwer Academic.
- Korsgaard, Christine. 1996a. *Sources of Normativity*. Cambridge: Cambridge University Press.

- Korsgaard, Christine. 1996b. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Korsgaard, Christine. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Ludwig, Bernd. 2010. "Die 'consequente Denkungsart der speculativen Kritik'. Kants radikale Umgestaltung seiner Freiheitslehre im Jahre 1786 und die Folgen für die Kritische Philosophie als Ganze." [The 'consequente Denkungsart der speculativen Kritik': Kant's Radical Transformation of his Doctrine of Freedom in the Year 1786 and Afterwards for the Critical Philosophy as a Whole.] *Deutsche Zeitschrift für Philosophie* 58: 595–628.
- Ludwig, Bernd. 2012. "Was weiss ich vom Ich? Kants Lehre vom Faktum der reinen praktischen Vernunft, seine Neufassung der Paralogismen und die verborgenen Fortschritte der Kritischen Metaphysik im Jahre 1786." [What do I know of the 'I'? Kant's Doctrine of the Fact of Pure Practical Reason, his New Version of Paralogisms and the Hidden Progress of Critical Metaphysics in the Year 1786.] In *Sind wir Bürger zweier Welten? Freiheit und moralische Verantwortung im transzendentalen Idealismus*, edited by Mario Brandhorst, Andree Hahmann and Bernd Ludwig, 155–194. Hamburg: Meiner.
- Ludwig, Bernd. 2015. "Die Kritik der reinen Vernunft hat die Wirklichkeit der Freiheit nicht bewiesen, ja nicht einmal deren Möglichkeit: Über die folgenreiche Fehlinterpretation eines Absatzes in der Kritik der reinen Vernunft." [The Critique of Pure Reason has not proved the actuality of freedom, nor even its possibility: On the Following Misinterpretation of a Paragraph in the Critique of Pure Reason.] *Kant-Studien* 106: 398–417.
- Łuków, Paweł. 1993. "The Fact of Reason: Kant's Passage to Ordinary Moral Knowledge." *Kant-Studien* 84: 204–221.
- McCarthy, Michael H. 1982. "Kant's Rejection of the Argument of *Groundwork* III." *Kant-Studien* 73: 169–190.
- Mellin, Georg Samuel Albert. 1800. *Encyclopädische Wörterbuch der Kritischen Philosophie* [*Encyclopedic Dictionary of Critical Philosophy*]. Leipzig: Frommann.
- Michaelis, Christian Friedrich. 1796–97. *Ueber die sittliche natur und bestimmung des menschen: ein versuch zur erläuterung über I. Kants Kritik der praktischen vernunft* [On the Moral Nature and Vocation of Human Beings: An Attempt at an Explanation of Kant's Critique of Practical Reason]. Leipzig: J.G. Beigang.
- Nagel, Thomas. 1970. *The Possibility of Altruism*. Princeton: Princeton University Press.
- Neiman, Susan. 1994. *The Unity of Reason: Rereading Kant*. New York, NY: Oxford University Press.
- O'Neill, Onora. 1989. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press.
- Paton, Herbert. 1947. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. Pennsylvania: University of Pennsylvania Press.
- Pistorius, Herman Andreas. 1786a. "Rezension der Grundlegung zur Metaphysik der Sitten." [Review of the *Groundwork for the Metaphysics of Morals*.] In *Materialien zu Kants Kritik der praktischen Vernunft (1775)*, edited by Rüdiger Bittner and Konrad Cramer. Frankfurt: Suhrkamp.
- Pistorius, Herman Andreas. 1786b. "Rezension von J. Schulzes Erläuterungen zur Kritik der reinen Vernunft." [Review of Schulz's Exposition of the Critique of Pure Reason.] In *Kants vergessener Rezensent. Die Kritik der theoretischen und praktischen Philosophie Kants in fünf frühen Rezensionen von Hermann Andreas Pistorius (2007)*, edited by Bernward Gesang. Hamburg: Meiner.

- Puls, Heiko. 2011. "Freiheit als Unabhängigkeit von bloß subjektiv bestimmenden Ursachen – Kants Auflösung des Zirkelverdachts im dritten Abschnitt der *Grundlegung zur Metaphysik der Sitten*." [Freedom as Independence from Merely Subjective Causes – Kant's Solution to the Threat of a Circle in the Third Section of the Groundwork for the Metaphysics of Morals.] In *Zeitschrift für Philosophische Forschung* 65: 534–562.
- Puls, Heiko. 2014. "Quo errat demonstrator – warum es in der Grundlegung eine Faktum-These gibt. Drei Argumente gegen Dieter Schöneckers Interpretation." [Quo errat demonstrator – Why there is a Fact-thesis in the Groundwork: Three Arguments Against Dieter Schönecker's Interpretation.] In *Kants Rechtfertigung des Sittengesetzes in Grundlegung III: Deduktion Oder Faktum?*, edited by Heiko Puls, 15–32. Berlin: Walter de Gruyter.
- Rauscher, Frederick. 2009. "Freedom and Reason in Groundwork III." In *Kant's Groundwork of the Metaphysics of Morals: A Critical Guide*, edited by Jens Timmermann, 203–223. Cambridge: Cambridge University Press.
- Rauscher, Frederick. 2015. *Naturalism and Realism in Kant's Ethics*. Cambridge: Cambridge University Press.
- Rawls, John. 2000. In *Lectures on the History of Moral Philosophy*, edited by Barbara Herman. Cambridge, MA: Harvard University Press.
- Reath, Andrews. 2012. "Kant's Moral Philosophy." In *The Oxford Handbook on the History of Ethics*, edited by Roger Crisp, 443–464. Oxford: Oxford University Press.
- Reinhold, Karl. 1786/87. "Briefe über die Kantische Philosophie." [Letters on Kantian Philosophy]. *Der Teutsche Merkur* 3: 99–127.
- Ross, William David. 1954. *Kant's Ethical Theory: A Commentary on the Grundlegung Zur Metaphysik Der Sitten*. Oxford: Clarendon Press.
- Saunders, Joe. 2014. "Kant, Rational Psychology and Practical Reason." *Kant Yearbook* 6: 115–135.
- Schönecker, Dieter. 1999. *Kant: Grundlegung III. Die Deduktion des kategorischen Imperativs [Kant: Groundwork III. The Deduction of the Categorical Imperative]*. Freiburg/München: Karl Alber-Verlag.
- Schönecker, Dieter. 2006. "How is a Categorical Imperative Possible?" In *Groundwork for the Metaphysics of Morals*, edited by Christoph Horn and Dieter Schönecker, 301–323. Berlin: De Gruyter.
- Schönecker, Dieter. 2013. "Kant's Moral Intuitionism: The Fact of Reason and Moral Predispositions." *Kant Studies Online* 1: 1–38.
- Schönecker, Dieter. 2014. "Warum es in der Grundlegung keine Faktum-These gibt. Drei Argumente." [Why there is No Fact-thesis in the Groundwork: Three Arguments.] In *Kants Rechtfertigung des Sittengesetzes in Grundlegung III: Deduktion Oder Faktum?*, edited by Heiko Puls, 1–15. Berlin: Walter de Gruyter.
- Schönecker, Dieter, ed. 2015. *Kants Begründung von Freiheit und Moral in 'Grundlegung III': Neue Interpretationen [Kant's Grounding of Freedom and Morality in Groundwork III: New Interpretations]*. Münster: Mentis.
- Stattler, Benedict. 1788. *Anti-Kant, Bd. 1–2 [Anti-Kant, Vols. 1–2]*. München: Lentner.
- Stern, Robert. 2012. *Understanding Moral Obligation: Kant, Hegel, Kierkegaard*. Cambridge: Cambridge University Press.
- Stern, Robert. 2015. *Kantian Ethics: Value, Agency, and Obligation*. Oxford: Oxford University Press.
- Sticker, Martin. 2015. "Educating the Common Agent: Kant on the Varieties of Moral Education." *Archiv für Geschichte der Philosophie* 97: 358–387.
- Sussman, David. 2008. "From Deduction to Deed: Kant's Grounding of the Moral Law." *Kantian Review* 13: 52–81.



- Timmermann, Jens. 2007. *Kant's Groundwork of the Metaphysics of Morals: A Commentary*. Cambridge: Cambridge University Press.
- Timmermann, Jens. 2010. "Reversal or Retreat? Kant's Deductions of Freedom and Morality." In *Kant's Critique of Practical Reason: A Critical Guide*, edited by Jens Timmermann and Andrews Reath, 73–89. Cambridge: Cambridge University Press.
- Tenenbaum, Sergio. 2012. "The Idea of Freedom and Moral Cognition in *Groundwork III*." *Philosophy and Phenomenological Research* 84: 555–589.
- Tittel, Gottlob August. 1786. *Ueber Herrn Kant's Moralreform [On Kant's Reform of Morals]*. Frankfurt: Pfähler.
- Velleman, David. 1989. *Practical Reflection*. Princeton: Princeton University Press.
- Velleman, David. 2009. *How We Get Along*. Cambridge: Cambridge University Press.
- Ware, Owen. 2014. "Rethinking Kant's Fact of Reason." *Philosophers' Imprint* 32: 1–21.
- Ware, Owen. 2016. "Skepticism in Kant's *Groundwork*." *European Journal of Philosophy* 24: 375–396.
- Williams, Terence Charles. 1968. *The Concept of the Categorical Imperative: A Study of the Place of the Categorical Imperative in Kant's Ethical Theory*. Oxford: Oxford University Press.
- Willaschek, Marcus. 1992. *Praktische Vernunft: Handlungstheorie und Moralbegründung bei Kant [Practical Reason: Action Theory and Moral Grounding in Kant]*. Weimar: Metzler.
- Wolff, Michael. 2009. "Warum das Faktum der Vernunft ein Faktum ist. Auflösung einiger Verständnisschwierigkeiten in Kants Grundlegung der Moral." [Why the Fact of Reason is a Fact: A Solution to Some Difficulties in Understanding Kant's *Groundwork of Morals*.] *Deutsche Zeitschrift für Philosophie* 57: 511–549.
- Wood, Allen W. 2011. "Preface and Introduction (3–16)." In *Immanuel Kant: Kritik der praktischen Vernunft*, edited by Otfried Höffe, 21–35. Berlin: Akademie Verlag.